BOUNDED RATIONALITY AND STATISTICAL PHYSICS

David H. Wolpert

NASA Ames Research Center

collectives.stanford.edu,

http://ic.arc.nasa.gov/people/dhw/

NASA-ARC-03-097

THE GOLDEN RULE

DO NOT:

Consider a variable x,
that optimizes a function

INSTEAD:

Consider a distribution over x,
that optimizes an expectation value

PROBABILITY COLLECTIVES (PC)

The golden rule gives an underlying language - *Probability Collectives* - for translating between many fields:

- **i) Bounded rational game theory**
- **ii) Statistical physics (mean field theory)**
- **iii) Adaptive control**
- **iv) Optimization, constrained or not, over *any* measurable space**
- **v) Reinforcement learning**
- **vi) Sampling of distributions**

Especially suited to distributed applications

ROADMAP

1) *Review information theory*
2) *Show bounded rationality = statistical physics*
3) *What is distributed control, formally?*
4) *Optimal distributed control policy*
5) *How to find that policy in a distributed way*

REVIEW OF INFORMATION THEORY

1) Want a quantification of how "uncertain" you are that you will observe a value i generated from $P(i)$.

2) Require the uncertainty at seeing the IID pair (i, i') to equal the sum of the uncertainties for i and for i'

3) This forces the definition

$$\boxed{\text{uncertainty}(i) = -\ln[P(i)]}$$

REVIEW OF INFORMATION THEORY - 2

4) So expected uncertainty is the *Shannon entropy*

$$\boxed{S(P) \equiv -\sum_i P(i) \ln[P(i)]}$$

- **Concave over P, infinite gradient at simplex border**

5) *Information* in P, I(P), is what's left after the uncertainty is removed: $-S$(P).

6) This allows us to formalize Occam's razor:

***Maxent*: Given $\{E_P(g_i) = 0\}$, "most plausible" P is the P consistent with $\{E_P(g_i) = 0\}$ having minimal I(P)**

ROADMAP

1) *Review information theory*

2) *Show bounded rationality = statistical physics*

3) *What is distributed control, formally?*

4) *Optimal distributed control policy*

5) *How to find that policy in a distributed way*

IN THE REAL WORLD, EVERYONE IS BOUNDED RATIONAL

- **Real players (human or otherwise) are bounded rational, due to limited computational power if nothing else.**

- **Previous attempts to address this are mostly ad hoc models of (human) players**

  - Underlying problem of arbitrariness of those models.

- **Science: Without information concerning a system, you cannot infer anything concerning it. So ...**

Inference of players' strategies *must* be based on observed/provided information.

COMBINING INFORMATION THEORY AND GAME THEORY

- **Say our information is $P_{(i)}$, the strategies of players other than i, and i's expected cost.**

  Then the minimum information principle says it is "most conservative" to infer that P_i minimizes

$$L_i(P) = \beta E(h_i) - S(P)$$

  where S is the Shannon entropy, and β a constant.

- **Alternative: If information is the entropy of i's mixed strategy, predict that P_i minimizes i's expected cost:**

Again, P_i minimizes $L_i(P)$

QUANTIFYING BOUNDED RATIONALITY

- **At Nash equilibrium, each P_i separately minimizes**

$$\boxed{E(h_i) = \int dz\, h_i(z) \prod_j P_j(z_j)}$$

- **Allow broader class of goals (*Lagrangians*) for the players**

Example

i) **Each P_i separately minimizes the Lagrangian**

$$\boxed{L_i(P) = \beta E(h_i) - S(P)}$$

for some appropriate function S (e.g., entropy ...)

ii) **$\beta < \infty$ is bounded rationality**

BOUNDED RATIONALITY AND COST OF COMPUTATION

- **Choose $\mathcal{S}(q) = \sum_i \int dz_i \mathcal{S}_i(P_i(z_i))$ (e.g., entropy).**

  Then bounded rationality is identical to conventional, full rationality — every player wants to minimize expected cost. *Only now there is a new cost function*:

$$f_i(z, P_i) \;=\; \beta h_i(z) \;-\; \mathcal{S}_i(P_i(z_i)) \,/\, P_i(z_i)$$

$-\mathcal{S}_i(P_i(z_i)) / P_i(z_i)$ measures the computational cost to player i for calculating $P_i(z_i)$

COMBINING GAME THEORY AND STATISTICAL PHYSICS

- **Jaynes showed that all statistical physics ensembles arise from minimizing**

  $$L_i(P) = \beta E(h_i) - S(P),$$

  with S the Shannon entropy

- **Mean field theory arises if P is a product distribution; bounded rational game theory = mean field theory (!)**

Much of the mathematics of statistical physics can be applied to bounded rational game theory

EXAMPLE: GAMES WITH VARIABLE NUMBERS OF PLAYERS

1) **The Grand Canonical Ensemble (GCE) of statistical physics models systems where the number of particles of various types varies stochastically.**

2) Use the underlying language, Probability Collectives, - which here is just Jaynesian inference - to translate the GCE into a game in which the number of players of various types can vary stochastically.

***Intuition*: Players with "types" = particles with properties**

GAMES WITH VARIABLE NUMBERS OF PLAYERS - 2

Example 1 (microeconomics):

i) A set of bounded rational companies,
ii) with cost functions given by market valuations,
iii) each of which must decide how many employees of various types to have.

Example 2 (evolutionary game theory):

i) A set of species,
ii) with cost functions given by fractions of total resources they consume,
iii) each of which must “decide” how many phenotypes of various types to express.

ROADMAP

1) *Review information theory*

2) *Show bounded rationality = statistical physics*

3) *What is distributed control, formally?*

4) *Optimal distributed control policy*

5) *How to find that policy in a distributed way*

DISTRIBUTED ADAPTIVE CONTROL

1) ***Control of routers in a network.***
2) ***Control of robots working together to construct a spacestation.***
3) ***Control of flaplets on an aircraft wing.***
4) ***Control of signals to human teams performing a joint task.***
5) ***Control of variables in a parallel computer algorithm to optimize a function.***

Must be adaptive (i.e., not wed to a system model) to

i) ***Avoid brittleness;***

ii) ***Scale well;***

iii) ***Be fault-tolerant;***

iv) ***Be widely applicable, with minimal (or even no) hand-tuning.***

WHAT IS DISTRIBUTED CONTROL?

1) ***A set of N agents: Joint move $x = (x_1, x_2, ..., x_N)$***

2) ***Since they are distributed, their joint probability is a product distribution:***

$$\boxed{q(x) = \prod_i q_i(x_i)}$$

- ***This definition of distributed agents is adopted from (extensive form) noncooperative game theory.***

3) ***Distributed control is a common payoff game - a bounded rational (statistical physics) one.***

EXAMPLE: KSAT

- $x = \{0, 1\}^N$

- **A set of many disjunctions, "clauses", each involving K bits.**
  E.g., $(x_2 \vee x_6 \vee \sim x_7)$ is a clause for K = 3

- **Goal: Find a bit-string x that simultaneously satisfies all clauses. G(x) is #violated clauses.**

- **For us, this goal becomes: find mixed strategy $q(x) = \prod_i q_i(x_i)$ tightly centered about such an x.**

The canonical computationally difficult problem

ITERATIVE DISTRIBUTED CONTROL

1) **s is current uncertainty of what x to pick, i.e., uncertainty of where $q(x)$ is concentrated.**

- **Early in the control process, high uncertainty.**

2) Find q minimizing $E_q(G)$ while consistent with s.

3) Reduce s. Return to (2).

4) Stop at mixed strategy q with good (low) $E_q(G)$.

Can do (2)→ (3) without ever explicitly specifying s

ITERATIVE DISTRIBUTED CONTROL - 2

1) **The central step is to "find the q that has lowest $E_q(G)$ while consistent with $S(q) = s$".**

2) **So we must find the critical point of the Lagrangian**

$$\boxed{L(q, T) \;=\; E_q(G) + T[s - S(q)]} \;,$$

i,e., find the q and T such that $\partial L/\partial q = \partial L/\partial T = 0$

- **Deep connections with statistical physics (*L* is "free energy" in mean-field theory), economics**

3) **Then we reduce *s*; repeat (find next critical point).**

EXAMPLE: KSAT

1) $S(q) = -\sum_i [b_i \ln(b_i) + (1 - b_i) \ln(1 - b_i)]$

where b_i is $q_i(x_i = \text{TRUE})$

2) $E_q(G) = \sum_{clauses\ j,\ x} q(x)\, K_j(x)$

$= \sum_{clauses\ j,\ x,\ i} \prod_i q_i(x_i)\, K_j(x)$

where $K_j(x) = 1$ iff x violates clause j

Our algorithm: i) Find q minimizing $E_q(G) - TS(q)$;

ii) Lower T and return to (i).

ROADMAP

1) ***Review information theory***

2) ***Show bounded rationality = statistical physics***

3) ***What is distributed control, formally?***

4) ***Optimal distributed control policy***

5) ***How to find that policy in a distributed way***

DISTRIBUTED SEARCH FOR q

So control reduces to finding q such that $\partial L/\partial q = 0$

1) **Since the agents make their moves in a distributed way, that q is a product distribution.**

2) **But they must also find that q in a distributed way.**

3) **There are two cases to consider:**
   i) Know functional form of G.
   ii) Don't know functional form of G - must sample.

MINIMIZING L(q) VIA GRADIENT DESCENT

1) Each i works to minimize $L(q_i, q_{(i)})$ using only partial information of the other agents' distribution, $q_{(i)}$.

2) The $q_i(x_i)$ component of $\nabla L(q)$, projected onto the space of allowed $q_i(x_i)$, is

$$\boxed{\begin{gathered}\beta\, E_{q_{(i)}}(G \mid x_i) \;+\; ln(q_i(x_i)) \\ - \\ \int dx'_i\, [\beta\, E_{q_{(i)}}(G \mid x_i) + ln(q_i(x'_i))\,]\end{gathered}}$$

• The subtracted term ensures q stays normalized

***GRADIENT DESCENT* - 2**

3) Each agent i knows its value of *ln*($q_i(x_i)$).

4) Each agent i knows the $E_{q_{(i)}}(G \mid x_i)$ terms.

Each agent knows how it should change

its q_i under gradient descent over L(q)

5) Gradient descent, even for categorical variables (!), and done in a distributed way.

6) Similarly the Hessian can readily be estimated (for Newton's method), etc.

EXAMPLE: KSAT

1) **Evaluate $\boxed{E_{q_{(i)}}(G \mid x_i)}$ - the expected number of violated clauses if bit i is in state x_i - for every i, x_i**

2) **In gradient descent, decrease each $q_i(x_i)$ by**

$$\boxed{\alpha[E_{q_{(i)}}(G \mid x_i) + T \ln[q_i(x_i)] - \mathrm{const}_j]}$$

where α is the stepsize, and const_j is an easy-to-evaluate normalization constant.

3) **We actually have a different T for each clause, and adaptively update all of them.**

ADAPTIVE DISTRIBUTED CONTROL

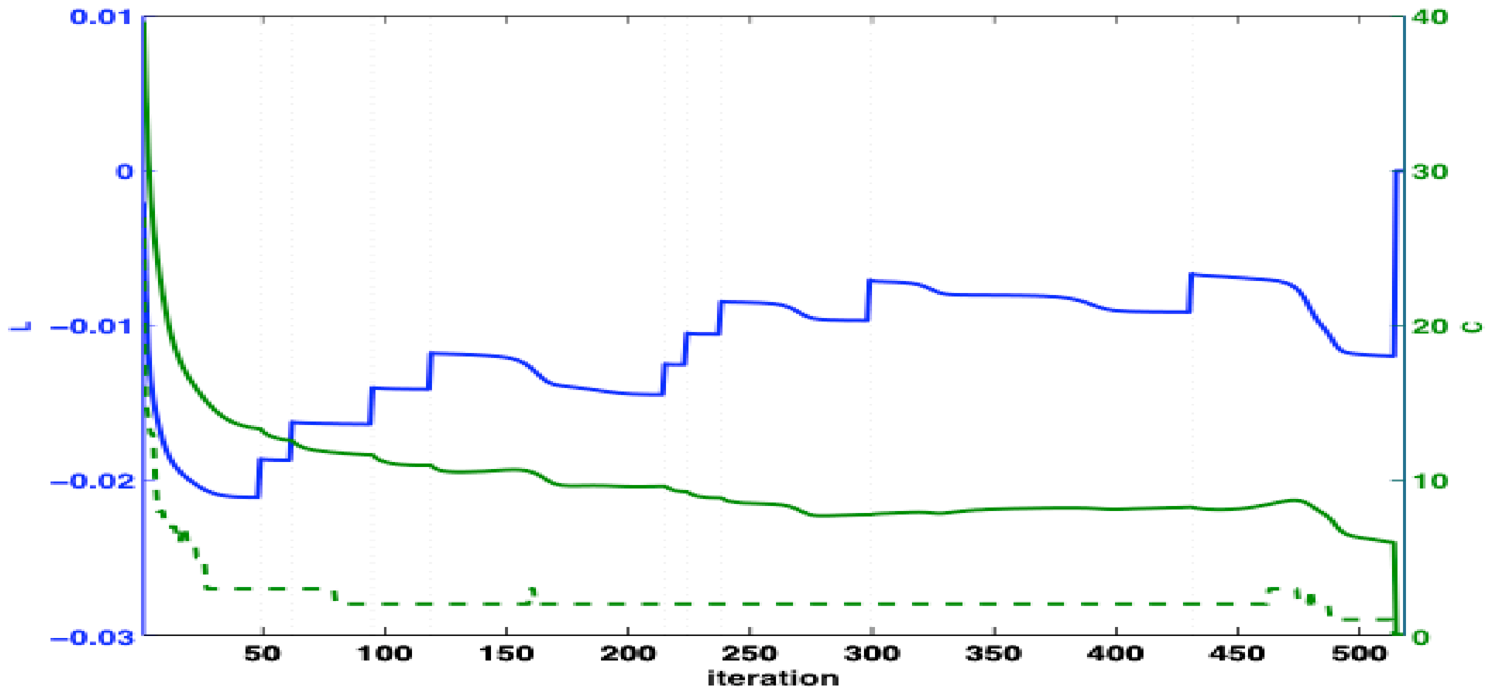
1) In *adaptive* control, don't know functional form of G(x). So use Monte Carlo:

- Sample G(x) repeatedly according to q;
- Each i independently estimates $E_{q_{(i)}}(G \mid x_i)$ for all its moves x_i;
- Only 1 MC process, no matter how many agents

So each q_i can adaptively estimate its update

EXAMPLE: KSAT

i) **Top plot is Lagrangian value vs. iteration;**
ii) **Middle plot is average (under q) number of constraint violations;**
iii) **Bottom plot is mode (under q) number of constraint violations.**



CONCLUSION

1) *Information theory - statistical inference - shows how to quantify bounded rationality*

2) *The same mathematics underlies stastistical physics; the two are identical.*

3) *That mathematics also underlies adaptive distributed control; all three are identical.*